PHONEMIC VOWEL LENGTH AND PROSODY IN AUSTRALIAN ENGLISH

Janet Fletcher *, Jonathan Harrington** and John Hajek*

*Department of Linguistics, University of Melbourne and **Speech Hearing and Language Research Centre, Macquarie University

ABSTRACT

Previous acoustic studies of Australian vowels note the similarity of formant patterns in the vowels /a:/ and /ʌ/. Most linguists and phoneticians describe this as an example of a vowel quantity versus vowel quality contrast. Two experiments were conducted to examine the acoustic and articulatory characteristics (i.e. jaw displacement) of these vowels in accented and unaccented environments. While significant formant and jaw height differences were observed in many instances, it appears these patterns can be interpreted as examples of articulatory undershoot, as opposed to inherent vowel quality differences.

INTRODUCTION

Vowel length is contrastive in many languages. Languages as typologically diverse as Swedish, Estonian, Tonkawa, and Kayardild contrast long and short vowels. Duration is also considered by phoneticians and phonologists to be an integral part of the tense/lax vowel distinction in many varieties of English , including Australian English. However, previous acoustic studies of Australian English vowels by Bernard (1987) and Cochrane (1970) note the acoustic similarity of the vowel nuclei in words like "part " and "putt" , and claim that this vowel pair constitutes an example of a relatively pure phonemic length contrast. Bernard's study encompassed 170 adult speakers of "broad, general and cultivated" varieties of Australian English. "Cultivated " speakers produced mean Formant 1 and 2 values of /a:/ 745 1325 and /ʌ/ 740 1430 whereas "broad" speakers produced values of 745 1325 and 755 and 1370, respectively. These formant values contrast with those cited by Fry (1979) for British English RP - /a:/ 680 1100 and /ʌ/ 720,1240 - a dialect of English in which the two vowels represent a vowel quality as well as vowel length distinction. F1 and F2 are lower in the long/tense vowel than in the short/lax vowel. In Australian English, by contrast the long vowel is much more centralised than in British English, and the formant pattern is reversed.

There are few known experimental articulatory studies of phonemic vowel length. Lindblom (1967) in a pioneering investigation, examined lip and jaw articulation in short and long vowels in Swedish. In most cases the degree of jaw lowering, relative to acoustic vowel duration was not significantly different in long and short vowels, although it appeared from his data, that the long open vowel /ɑ/ was realised with a slightly lower displacement (1-3 mm) than its short counterpart /a/. These results are compatible with general intuition amongst phoneticians and phonologists that vowel quantity differences are often perceived as slight quality differences. It is not clear, however, how to determine the boundary between a phonemic length contrast and a qualitative contrast, such as the tense/lax distinction. Furthermore, it is also apparent that there may be more than one articulatory strategy to realise an underlying phonological contrast. In a recent electro-microbeam study of tense/lax vowels in American English, Johnson et al. (1993) discuss three different strategies to make the tense/lax distinction , one in which tongue height is varied independently of the jaw to realise the two vowels, another in which the jaw is directly coupled with tongue height, and finally another strategy in which the jaw and the tongue coordinate in entirely opposite directions. Degree of openness seems more or less directly correlated with jaw height and the level of F1, whereas oppositions within the tense/lax category are realised as variations in tongue elevation.

The aims of current study were to investigate the acoustic and articulatory characteristics of the /a:/ /ʌ/ contrast in Australian English. The Bernard study comprised 3 repetitions of the words 'hud' and 'hard' per speaker. The first experiment was designed to investigate Bernard and Cochrane's findings in a more general corpus of Australian English. The influence of stress was also considered in this experiment. It is well known that degree of prosodic prominence influences formant structure and duration patterns of syllables. The second experiment was designed to examine the articulatory correlates of the vowel contrast in the light of earlier kinematic investigations of vowel articulation and other duration-influencing factors such as degree of prominence (eg. Edwards et al., 1991, Harrington et al., in press). Vowel lengthening due to accentuation appears to be best modelled as the result of decreased articulatory overlap between successive opening and closing lip/jaw gestures. As a result, accented vowels are associated with bigger, longer articulatory gestures, than unaccented vowels. Phonemic vowel length differences may well be the product of a similar articulatory strategy. However, an accompanying acoustic study was also carried out, in order to counter some of the potential speaker-specific strategies, highlighted by the Johnson et al. study.

EXPERIMENT 1

Subjects and Materials

Speech data from the Australian National Database of Spoken Language (Millar et al. 1990) formed the corpus for this experiment. Data from 5 speakers reading 400 phonetically dense and balanced sentences were recorded and digitised at 20 Khz using ESPS Waves + running on Sun Work stations at the Speech Hearing and Language Research Centre Macquarie University. The data were annotated (see Croot et al. 1992 for further information on levels of labelling), following standard acoustic phonetic segmentation procedures (eg. Barry and Fourcin 1992).

The Mu+ database management system was used to retrieve and analyse all instances of the vowels /a:/ and /ʌ/ and their accompanying signal files. For each speaker , mean and standard deviation values of F1 and F2 were computed using Mu+ . A 2-way ANOVA was performed to test the significance of any vowel/formant frequency interactions. Further analysis was performed on vowels according to degree of stress.

Results and Discussion

The results of the acoustic analysis are illustrated in Table I. A total of 123 instances of the vowel /a:/ and 203 instances of the vowel /ʌ/ were retrieved for each speaker . All speakers produced a significant overall vowel length contrast (JC: F=191.14, DB: F=242.43, DW: F=77.12, MB: F=193.89, ND: F=254.49). The rest of the results will be discussed according to stress category. Two speakers produced significantly different F1 and F2 values for the two vowels (JC: F=4.4, DW: F=15.34) although the effect was somewhat weaker for speaker JC. There were significant interactions between vowel identity and F1 and F2 (JC: F=67.4, DB: F=44.73, DW: F=8.17, MB: F=4.37, ND: F=6.1) across the corpus. In unstressed tokens, these differences were no longer significant for two speakers (DW and MB). However the remaining speakers showed varying degrees of interaction between vowel identity and F1 and F2 values (JC: F=18.69, DB: F=8.13, ND: F=17.97).

With respect to the influence of prosodic prominence on individual formants, stress consistently raised F1 in all speakers' data. These results are similar to those discussed by Summers (1988), for example , who found that F1 was raised in stressed low vowels in American English. Furthermore, similar effects of vowel and prominence category were noted (JC: F=14.62 x vowel, F=14.38 x stress; DW F=23.41; $p<0.0001$ x stress; MB; F=33.8 x vowel, 11.94 x prominence ; ND: 11.49 x vowel; F=11.26 x prominence). The effects of stress on F2 were less clear. F2 was signficantly lower in stressed vowels for 2 speakers (JC and ND, F's 42.92,5.9, respectively). Vowel identity had a significant influence on F2 values for 3 speakers (JC, DW, and MB - F's 42.9, 9.12, 8.5).

One possible interpretation of these results could be that the minor formant differences are not sufficient to produce a tense/lax vowel quality. Furthermore, formant differences between the short and long vowel could be the result of the same kind of articulatory strategy involved in accenting a full vowel. Edwards et al. (1991) and Harrington et al. (in press) have hypothesised that de-accenting, by contrast, can be modelled as undershoot. In other words, both vowels have the same underlying gestural target (in this case, jaw lowering or tongue elevation), however, in short vowels, the closing gesture begins earlier than it would in long vowels, resulting in modified formant frequencies as well as shorter acoustic durations. This was one of the issues addressed in Experiment 2.

		/a:/			/ʌ/		
		F1	*F2*	*Dur(ms)*	*F1*	*F2*	*Dur(ms)*
Speaker JC	*Stressed*	641	1251	164	570	1295	93
	Full Vowel	571	1262	152	525	1324	85
Speaker DB	*Stressed*	633	1364	163	575	1395	79
	Full Vowel	614	1395	134	565	1438	75
Speaker DW	*Stressed*	831	1336	176	751	1331	96
	Full Vowel	770	1378	161	730	1412	92
Speaker MB	*Stressed*	628	1278	155	569	1303	82
	Full Vowel	610	1275	131	545	1315	72
Speaker ND	*Stressed*	699	1267	168	680	1310	88
	Full Vowel	715	1295	139	625	1372	82

Table I. Mean Formant Frequency (Hz) and Duration (ms) values for /a/ and /ʌ/

EXPERIMENT 2

Subjects and Materials

Three speakers of general/educated Australian English produced 15 repetitions of the following sentence at two self-selected tempi; conversational and fast. For the purposes of the current study, only the conversational tempo utterances will be considered. The materials listed below were designed to elicit an accented (Sentences 1 and 2) and an unaccented production (Sentences 3 and 4) of the italicised words.

(1) Say *barb* naturally, not bub. (2) Say *bub* naturally, not barb.
(3) Say barb *naturally,* not slowly. (4) Say bub *naturally*, not slowly.

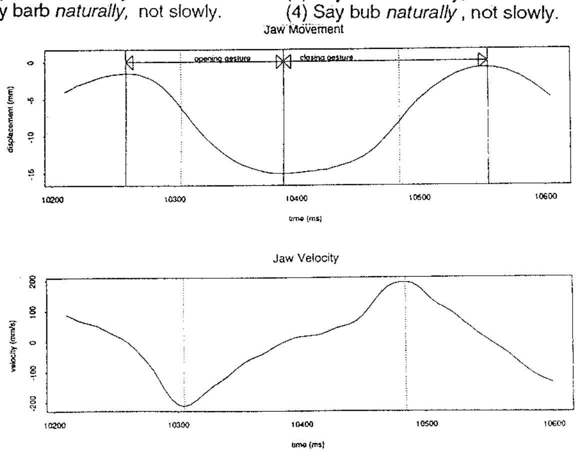

Figure 1. Movetrack trace representing vertical jaw position and instantaneous velocity for a repetition of "barb".

The three speakers were requested to recite the sentence as two prosodic phrases corresponding to the portions of the orthographic text divided by commas. Recordings of the speech waveform and six articulatory channels (lower lip, upper lip, jaw, all in the horizontal and vertical dimension) were made using the Movetrack Articulatory tracking system interfaced to Waves+. Acoustic segmentation and labelling of the two words "barb" and "bub" was performed, and the acoustic and kinematic data read into the mu+ speech database analysis system (Harrington et al., 1993). Acoustic measurements of vowel duration and formants 1 and 2 were obtained . We chose to analyse jaw movement in the vertical dimension because jaw height is a known indicator of degree of openness of a vowel (Lindblom, 1967). For position of the jaw, (fig 1) peaks of the movement trace correspond to points of maximum closure for the bilabial consonant, and valleys correspond to points of maximum lowering or opening relating to the vowel gesture (/a:/ or /ʌ/) . Measurements of maximum jaw lowering were obtained for each experimental token.

Results and Discussion

Table II shows the results of the acoustic and articulatory analyses of /a:/ and /ʌ/. Figure 2 shows mean Formant 1 and 2 frequencies for 2 speakers. All speakers produced significantly different durations for the two vowels in the accented and unaccented conditions (KC: Accented, $F(2,26)= 128.13; p<0.0001$; Unaccented ,$F(2,26)= 118$; $p<0.0001$;JF: Accented,$F(2,28)=118.75$; $p<0.0001$ - Unaccented, $F(2,28)=37.84$; CR : Accented, $F(2,29)= 130.12$; $p<0.0001$ - Unaccented, $F(2,29)=35.07$). In the accented tokens, only KC produced significant F1 and F2 differences for the short versus long vowel (KC: $F= 11.73$; $p<0.001$). Simple main effects analysis of the effects of vowel identity on individual formants shows that F1 is slightly although consistently lower in the short vowel for this speaker (KC: $F=12.35$; $p<0.001$). In unaccented tokens, however, formant values are not significantly different across the corpus.

		/a:/				/ʌ/			
		F1	*F2*	*D (ms)*	*Jaw Disp(mm)*	*F1*	*F2*	*D(ms)*	*Jaw Disp(mm)*
CR	*Accented*	816	1382	189	16.8	793	1354	112	13.22
	Unaccented	796	1459	136	11.5	758	1416	103	10.12
JF	*Accented*	824	1416	180	11.4	818	1395	114	9.0
	Unaccented	853	1418	97	7.7	822	1385	85	7.2
KC	*Accented*	840	1350	241	12.4	750	1482	134	10.9
	Unaccented	798	1526	148	9.9	776	1501	105	8.9

Table II . Mean values of Formant 1 and 2 frequencies (hz) , acoustic duration (ms) and maximum jaw opening for vowels /a:/ and /ʌ/.

All three speakers exhibit slightly lower jaw positions for the long versus short vowel in accented tokens. These differences are significant at the .001 level, although they are very small, ie. from 1-3mm. In unaccented tokens, only two of the three speakers show a slightly significant lower jaw position in longer vowels (CR : $F=4.25$ and KC: $F=6.16$; $p<0.0001$) with differences in the range of .7 - 1 mm. These results are similar to those found for the low vowel phonemic length distinction in Swedish (Lindblom, 1967). On the basis of the acoustic and articulatory data in this experiment, it appears that two of the three speakers are producing a relatively straightforward phonemic vowel quantity as opposed to vowel quality contrast. It is not clear, however, whether the formant differences produced by KC are sufficient to constitute a vowel quality contrast.

Whilst F1 is clearly correlated with extent of jaw lowering in these data, it seems there is a more complex relationship between the two parameters. With respect to speaker-specific patterns, the significant 100 hz difference in F1 for KC's accented productions of /a:/ and /ʌ/ corresponds to a 1.5mm difference in jaw opening, suggesting potential interplay between pharyngeal enlargement due to some combination of jaw lowering, tongue elevation, and larynx lowering to produce the acoustic difference. A converse strategy might have been adopted by speaker CR, who produced the long vowel with a 3.6mm difference in jaw height, which translated in the acoustic domain as an insignificant 23 hz formant difference. These different articulatory strategies fit well with Johnson et al. 's analysis of the articulatory correlates of the tense/lax distinction.

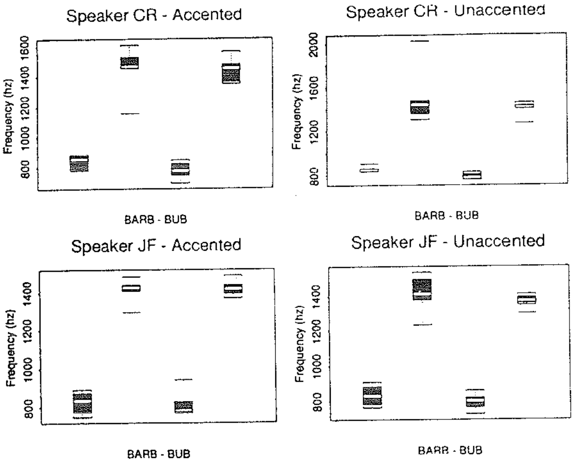

Figure 2. Mean F1 and F2 values for two speakers' productions of /a:/ and /ʌ/.

CONCLUSION

The results of these two experiments confirm, in part, Bernard and Cochrane's earlier findings. The /a:/ v. /ʌ/ contrast is realised primarily as a duration contrast. This is clearly observed in Experiment 2. The qualitative formant frequency patterns of each vowel pattern quite differently from those observed for the same phonemic contrast in RP. In Experiment 1, the data showed a greater amount of variation than observed by Bernard. This was due primarily to the different nature of the data studied (general corpus, versus isolated tokens) and the fact that the data were not pooled across speakers . The range of variation in experiment 1 might also be explained by dialectal differences within Australian English. The two speakers that produced consistently different F1 and F2 frequencies for the vowel pair could be classified as speakers near the top range of the general/cultivated dialect continuum. Further analysis of the jaw data in Experiment 2 is currently underway to test the hypothesis that short vowels are merely "undershot" long vowels.

Acknowledgments: This research was supported by the Australian Research Council.

REFERENCES

Barry, W.J. and Fourcin, A. (1992) "Levels of Labelling", Computer, Speech and Language, 6, 1-14

Bernard, J. (1989) "Quantitative aspects of the sounds of Australian English" in P. Collins and D. Blair (eds) *Australian English: the language of a new society*. St. Lucia: University of Queensland Press. p.

187-204.

Cochrane, G.R. (1970) "Some vowel durations in Australian English". Phonetica, 22, 240-250.

Croot, K., Fletcher, J. and Harrington, J. (1992). "Levels of segmentation and labelling in the Australian Database of Spoken Language", Proc. 4th Australian Internat. Conf. Speech Science and Speech Technology., 86-91.

Edwards, J., Beckman, M. and Fletcher, J. (1991). "The articulatory kinematics of final lengthening". Journal of the Acoustical Society of America, 89, 369-382.

Fletcher, J. and McVeigh, A. (1993). "Segment and syllable duration in Australian English". Speech Communication 13, 355-365.

Harrington, J., Cassidy, S., Fletcher, J. and McVeigh, A. (1993) "The mu+ system for corpus based speech research. Computer Speech and Language, 7, 305-331.

Harrington, J., Fletcher, J. and Roberts, C. (in press). "Coarticulation and the accented/unaccented distinction: evidence from jaw movement data". Journal of Phonetics.

Johnson, K., Ladefoged, P. and Lindau, M. (1993). "Individual differences in vowel production". Journal of the Acoustical Society of America, 94, 701-714.

Lindblom, B.E.F. (1963). " Spectrographic study of vowel reduction". Journal of the Acoustical Society of America, 35, 1773-1781.

Lindblom, B.E.F. (1967) "Vowel duration and a model of lip mandible coordination" STL- QPSR, 4, 1-29.

Millar, J., Dermody, P., Harrington, J. and Vonwiller, J. (1990). "A national database of spoken language: concept, design, and implementation". Proceedings of the International Conference on Spoken Language Processing (ICSLP-90). Kobe, Japan.

Summers, W.V. (1988). "F1 structure provides information from final consonant voicing". Journal of the Acoustical Society of America, 84, 485-492.